



SEQUENCE LISTING

<110> Madison, Wisconsin

Office of Patent and Trademark Office

Yeh, Juinn-Chern

<120> NUCLEIC ACID MOLECULES ENCODING SERINE PROTEASE 16, THE
ENCODED PROTEINS AND METHODS BASED THEREON

<130> 24745-1625

<140> 10/612,466

<141> 2003-07-01

<150> 60/394,347

<151> 2002-07-02

<160> 22

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3147

<212> DNA

<213> Homo Sapien

<220>

<221> CDS

<222> (23) ... (2589)

<223> Nucleotide sequence encoding MTSP1

<300>

<308> GenBank #AR081724

<309> 2000-08-31

<400> 1

tcaagagcgg cctcgggta cc atg ggg agc gat cg^g gcc cgc a^g ggc gga 52
Met Gly Ser Asp Arg Ala Arg Lys Gly Gly
1 5 10gg^g ggc cc^g a^a g^a c^c t^t g^g c^g g^g a^a g^a c^c t^t c^t a^a c^c t^t c^g c^a 100
Gly Gly Pro Lys Asp Phe Gly Ala Gly Leu Lys Tyr Asn Ser Arg His
15 20 25gag aaa gt^g aat g^gc tt^g gag gaa g^gc gt^g gag tt^c ct^g c^{ca} g^{tc} a^{ac} 148
Glu Lys Val Asn Gly Leu Glu Gly Val Phe Leu Pro Val Asn
30 35 40aac g^{tc} a^a g^a gt^g gaa a^a g^a c^{at} g^{gc} c^{cg} g^{gg} c^{gc} t^{gg} g^{tg} g^{tg} ct^g 196
Asn Val Lys Lys Val Glu Lys His Gly Pro Gly Arg Trp Val Val Leu
45 50 55g^{ca} g^{cc} g^{tg} c^{tg} a^{tc} g^{gc} c^{tc} c^{tc} t^{tg} g^{tc} t^{tg} c^{tg} g^{gg} a^{tc} g^{gc} t^{tc} 244
Ala Ala Val Leu Ile Gly Leu Leu Val Leu Leu Gly Ile Gly Phe
60 65 70ct^g g^{tg} t^{gg} c^{at} t^{tg} c^{ag} t^{ac} c^{gg} g^{ac} g^{tg} c^{gt} g^{tc} c^{ag} a^a g^{tc} t^{tc} 292
Leu Val Trp His Leu Gln Tyr Arg Asp Val Arg Val Gln Lys Val Phe
75 80 85 90

aat ggc tac atg agg atc aca aat gag aat ttt gtg gat gcc tac gag Asn Gly Tyr Met Arg Ile Thr Asn Glu Asn Phe Val Asp Ala Tyr Glu	95	100	105	340
aac tcc aac tcc act gag ttt gta agc ctg gcc agc aag gtg aag gac Asn Ser Asn Ser Thr Glu Phe Val Ser Leu Ala Ser Lys Val Lys Asp	110	115	120	388
gcg ctg aag ctg ctg tac agc gga gtc cca ttc ctg ggc ccc tac cac Ala Leu Lys Leu Leu Tyr Ser Gly Val Pro Phe Leu Gly Pro Tyr His	125	130	135	436
aag gag tcg gct gtg acg gcc ttc agc gag ggc agc gtc atc gcc tac Lys Glu Ser Ala Val Thr Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr	140	145	150	484
tac tgg tct gag ttc agc atc ccg cag cac ctg gtg gag gag gcc gag Tyr Trp Ser Glu Phe Ser Ile Pro Gln His Leu Val Glu Glu Ala Glu	155	160	165	532
cgc gtc atg gcc gag gag cgc gta gtc atg ctg ccc ccg cgg gcg cgc Arg Val Met Ala Glu Glu Arg Val Val Met Leu Pro Pro Arg Ala Arg	175	180	185	580
tcc ctg aag tcc ttt gtg gtc acc tca gtg gtg gct ttc ccc acg gac Ser Leu Lys Ser Phe Val Val Thr Ser Val Val Ala Phe Pro Thr Asp	190	195	200	628
tcc aaa aca gta cag agg acc cag gac aac agc tgc agc ttt ggc ctg Ser Lys Thr Val Gln Arg Thr Gln Asp Asn Ser Cys Ser Phe Gly Leu	205	210	215	676
cac gcc cgc ggt gtg gag ctg atg cgc ttc acc acg ccc ggc ttc cct His Ala Arg Gly Val Glu Leu Met Arg Phe Thr Thr Pro Gly Phe Pro	220	225	230	724
gac agc ccc tac ccc gct cat gcc cgc tgc cag tgg gcc ctg cgg ggg Asp Ser Pro Tyr Pro Ala His Ala Arg Cys Gln Trp Ala Leu Arg Gly	235	240	245	772
gac gcc gac tca gtg ctg agc ctc acc ttc cgc agc ttt gac ctt gcg Asp Ala Asp Ser Val Leu Ser Leu Thr Phe Arg Ser Phe Asp Leu Ala	255	260	265	820
tcc tgc gac gag cgc ggc agc gac ctg gtg acg gtg tac aac acc ctg Ser Cys Asp Glu Arg Gly Ser Asp Leu Val Thr Val Tyr Asn Thr Leu	270	275	280	868
agc ccc atg gag ccc cac gcc ctg gtg cag ttg tgt ggc acc tac cct Ser Pro Met Glu Pro His Ala Leu Val Gln Leu Cys Gly Thr Tyr Pro	285	290	295	916
ccc tcc tac aac ctg acc ttc cac tcc cag aac gtc ctg ctc atc Pro Ser Tyr Asn Leu Thr Phe His Ser Ser Gln Asn Val Leu Leu Ile	300	305	310	964
aca ctg ata acc aac act gag cgg cgg cat ccc ggc ttt gag gcc acc Thr Leu Ile Thr Asn Thr Glu Arg Arg His Pro Gly Phe Glu Ala Thr	315	320	325	1012
				330

ttc ttc cag ctg cct agg atg agc agc tgt gga ggc cgc tta cgt aaa Phe Phe Gln Leu Pro Arg Met Ser Ser Cys Gly Gly Arg Leu Arg Lys 335	340	345	1060
gcc cag ggg aca ttc aac agc ccc tac tac cca ggc cac tac cca ccc Ala Gln Gly Thr Phe Asn Ser Pro Tyr Tyr Pro Gly His Tyr Pro Pro 350	355	360	1108
aac att gac tgc aca tgg aac att gag gtg ccc aac aac cag cat gtg Asn Ile Asp Cys Thr Trp Asn Ile Glu Val Pro Asn Asn Gln His Val 365	370	375	1156
aag gtg agc ttc aaa ttc ttc tac ctg ctg gag ccc ggc gtg cct gcg Lys Val Ser Phe Lys Phe Tyr Leu Leu Glu Pro Gly Val Pro Ala 380	385	390	1204
ggc acc tgc ccc aag gac tac gtg gag atc aat ggg gag aaa tac tgc Gly Thr Cys Pro Lys Asp Tyr Val Glu Ile Asn Gly Glu Lys Tyr Cys 395	400	405	1252
gga gag agg tcc cag ttc gtc gtc acc agc aac agc aac aag atc aca Gly Glu Arg Ser Gln Phe Val Val Thr Ser Asn Ser Asn Lys Ile Thr 415	420	425	1300
gtt cgc ttc cac tca gat cag tcc tac acc gac acc ggc ttc tta gct Val Arg Phe His Ser Asp Gln Ser Tyr Thr Asp Thr Gly Phe Leu Ala 430	435	440	1348
gaa tac ctc tcc tac gac tcc agt gac cca tgc ccg ggg cag ttc acg Glu Tyr Leu Ser Tyr Asp Ser Ser Asp Pro Cys Pro Gln Phe Thr 445	450	455	1396
tgc cgc acg ggg cgg tgt atc cgg aag gag ctg cgc tgt gat ggc tgg Cys Arg Thr Gly Arg Cys Ile Arg Lys Glu Leu Arg Cys Asp Gly Trp 460	465	470	1444
gcc gac tgc acc gac cac agc gat gag ctc aac tgc agt tgc gac gcc Ala Asp Cys Thr Asp His Ser Asp Glu Leu Asn Cys Ser Cys Asp Ala 475	480	485	1492
ggc cac cag ttc acg tgc aag aac aag ttc tgc aag ccc ctc ttc tgg Gly His Gln Phe Thr Cys Lys Asn Lys Phe Cys Lys Pro Leu Phe Trp 495	500	505	1540
gtc tgc gac agt gtg aac gac tgc gga gac aac agc gac gag cag ggg Val Cys Asp Ser Val Asn Asp Cys Gly Asp Asn Ser Asp Glu Gln Gly 510	515	520	1588
tgc agt tgt ccg gcc cag acc ttc agg tgt tcc aat ggg aag tgc ctc Cys Ser Cys Pro Ala Gln Thr Phe Arg Cys Ser Asn Gly Lys Cys Leu 525	530	535	1636
tcg aaa agc cag cag tgc aat ggg aag gac gac tgt ggg gac ggg tcc Ser Lys Ser Gln Gln Cys Asn Gly Lys Asp Asp Cys Gly Asp Gly Ser 540	545	550	1684
gac gag gcc tcc tgc ccc aag gtg aac gtc gtc act tgt acc aaa cac Asp Glu Ala Ser Cys Pro Lys Val Asn Val Val Thr Cys Thr Lys His 555	560	565	1732
		570	

acc tac cgc tgc ctc aat ggg ctc tgc ttg agc aag ggc aac cct gag	1780
Thr Tyr Arg Cys Leu Asn Gly Leu Cys Leu Ser Lys Gly Asn Pro Glu	
575	580
585	
tgt gac ggg aag gag gac tgt agc gac ggc tca gat gag aag gac tgc	1828
Cys Asp Gly Lys Glu Asp Cys Ser Asp Gly Ser Asp Glu Lys Asp Cys	
590	595
600	
gac tgt ggg ctg cgg tca ttc acg aga cag gct cgt gtt gtt ggg ggc	1876
Asp Cys Gly Leu Arg Ser Phe Thr Arg Gln Ala Arg Val Val Gly Gly	
605	610
615	
acg gat gcg gat gag ggc gag tgg ccc tgg cag gta agc ctg cat gct	1924
Thr Asp Ala Asp Glu Gly Glu Trp Pro Trp Gln Val Ser Leu His Ala	
620	625
630	
ctg ggc cag ggc cac atc tgc ggt gct tcc ctc atc tct ccc aac tgg	1972
Leu Gly Gln Gly His Ile Cys Gly Ala Ser Leu Ile Ser Pro Asn Trp	
635	640
645	650
ctg gtc tct gcc gca cac tgc tac atc gat gac aga gga ttc agg tac	2020
Leu Val Ser Ala Ala His Cys Tyr Ile Asp Asp Arg Gly Phe Arg Tyr	
655	660
665	
tca gac ccc acg cag tgg acg gcc ttc ctg ggc ttg cac gac cag agc	2068
Ser Asp Pro Thr Gln Trp Thr Ala Phe Leu Gly Leu His Asp Gln Ser	
670	675
680	
cag cgc agc gcc cct ggg gtg cag gag cgc agg ctc aag cgc atc atc	2116
Gln Arg Ser Ala Pro Gly Val Gln Glu Arg Arg Leu Lys Arg Ile Ile	
685	690
695	
tcc cac ccc ttc aat gac ttc acc ttc gac tat gac atc gcg ctg	2164
Ser His Pro Phe Phe Asn Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu	
700	705
710	
ctg gag ctg gag aaa ccg gca gag tac agc tcc atg gtg cgg ccc atc	2212
Leu Glu Leu Glu Lys Pro Ala Glu Tyr Ser Ser Met Val Arg Pro Ile	
715	720
725	730
tgc ctg ccg gac gcc tcc cat gtc ttc cct gcc ggc aag gcc atc tgg	2260
Cys Leu Pro Asp Ala Ser His Val Phe Pro Ala Gly Lys Ala Ile Trp	
735	740
745	
gtc acg ggc tgg gga cac acc cag tat gga ggc act ggc gcg ctg atc	2308
Val Thr Gly Trp Gly His Thr Gln Tyr Gly Gly Thr Gly Ala Leu Ile	
750	755
760	
ctg caa aag ggt gag atc cgc gtc atc aac cag acc acc tgc gag aac	2356
Leu Gln Lys Gly Glu Ile Arg Val Ile Asn Gln Thr Thr Cys Glu Asn	
765	770
775	
ctc ctg ccg cag cag atc acg ccg cgc atg atg tgc gtg ggc ttc ctc	2404
Leu Leu Pro Gln Gln Ile Thr Pro Arg Met Met Cys Val Gly Phe Leu	
780	785
790	
agc ggc ggc gtg gac tcc tgc cag ggt gat tcc ggg gga ccc ctg tcc	2452
Ser Gly Gly Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser	
795	800
805	810

agc gtg gag gcg gat ggg cgg atc ttc cag gcc ggt gtg gtg agc tgg 2500
 Ser Val Glu Ala Asp Gly Arg Ile Phe Gln Ala Gly Val Val Ser Trp
 815 820 825
 gga gac ggc tgc gct cag agg aac aag cca ggc gtg tac aca agg ctc 2548
 Gly Asp Gly Cys Ala Gln Arg Asn Lys Pro Gly Val Tyr Thr Arg Leu
 830 835 840
 cct ctg ttt cgg gac tgg atc aaa gag aac act ggg gta ta ggggccccggg 2599
 Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr Gly Val
 845 850 855
 ccacccaaat gtgtacacct gcggggccac ccatacgcca ccccagtgtg cacgcctgca 2659
 ggctggagac tggaccgctg actgcaccag cgccccccaga acatacactg tgaactcaat 2719
 ctccagggtc ccaaattctgc ctagaaaacc tctcgcttcc tcagcctcca aagtggagct 2779
 gggaggtaga aggggaggac actgggtggtt ctactgaccc aactgggggc aaagggttga 2839
 agacacagcc tcccccccca gcccccaagct gggccgaggc gcgtttgtgt atatctgcct 2899
 cccctgtctg taaggagcac cgggaacggg gcttcggagc ctcctcagtg aaggtgggtgg 2959
 ggctgcccggg tctgggctgt gggcccttg ggccacgctc ttgaggaagc ccagggttcgg 3019
 aggaccctgg aaaacagacg ggtctgagac tgaaattgtt ttaccagctc ccagggtggg 3079
 cttcagtgtg tgtatgggtg taaatgggtt aaacaattta tttttttt aaaaaaaaaa 3139
 aaaaaaaaaa 3147

<210> 2
 <211> 855
 <212> PRT
 <213> Homo Sapien

<400> 2
 Met Gly Ser Asp Arg Ala Arg Lys Gly Gly Gly Pro Lys Asp Phe
 1 5 10 15
 Gly Ala Gly Leu Lys Tyr Asn Ser Arg His Glu Lys Val Asn Gly Leu
 20 25 30
 Glu Glu Gly Val Glu Phe Leu Pro Val Asn Asn Val Lys Lys Val Glu
 35 40 45
 Lys His Gly Pro Gly Arg Trp Val Val Leu Ala Ala Val Leu Ile Gly
 50 55 60
 Leu Leu Leu Val Leu Leu Gly Ile Gly Phe Leu Val Trp His Leu Gln
 65 70 75 80
 Tyr Arg Asp Val Arg Val Gln Lys Val Phe Asn Gly Tyr Met Arg Ile
 85 90 95
 Thr Asn Glu Asn Phe Val Asp Ala Tyr Glu Asn Ser Asn Ser Thr Glu
 100 105 110
 Phe Val Ser Leu Ala Ser Lys Val Lys Asp Ala Leu Lys Leu Leu Tyr
 115 120 125
 Ser Gly Val Pro Phe Leu Gly Pro Tyr His Lys Glu Ser Ala Val Thr
 130 135 140
 Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser
 145 150 155 160
 Ile Pro Gln His Leu Val Glu Glu Ala Glu Arg Val Met Ala Glu Glu
 165 170 175
 Arg Val Val Met Leu Pro Pro Arg Ala Arg Ser Leu Lys Ser Phe Val
 180 185 190
 Val Thr Ser Val Val Ala Phe Pro Thr Asp Ser Lys Thr Val Gln Arg
 195 200 205
 Thr Gln Asp Asn Ser Cys Ser Phe Gly Leu His Ala Arg Gly Val Glu
 210 215 220
 Leu Met Arg Phe Thr Thr Pro Gly Phe Pro Asp Ser Pro Tyr Pro Ala
 225 230 235 240
 His Ala Arg Cys Gln Trp Ala Leu Arg Gly Asp Ala Asp Ser Val Leu

	245	250	255
Ser Leu Thr Phe Arg Ser Phe Asp Leu Ala Ser Cys Asp Glu Arg Gly			
260	265	270	
Ser Asp Leu Val Thr Val Tyr Asn Thr Leu Ser Pro Met Glu Pro His			
275	280	285	
Ala Leu Val Gln Leu Cys Gly Thr Tyr Pro Pro Ser Tyr Asn Leu Thr			
290	295	300	
Phe His Ser Ser Gln Asn Val Leu Leu Ile Thr Leu Ile Thr Asn Thr			
305	310	315	320
Glu Arg Arg His Pro Gly Phe Glu Ala Thr Phe Phe Gln Leu Pro Arg			
325	330	335	
Met Ser Ser Cys Gly Gly Arg Leu Arg Lys Ala Gln Gly Thr Phe Asn			
340	345	350	
Ser Pro Tyr Tyr Pro Gly His Tyr Pro Pro Asn Ile Asp Cys Thr Trp			
355	360	365	
Asn Ile Glu Val Pro Asn Asn Gln His Val Lys Val Ser Phe Lys Phe			
370	375	380	
Phe Tyr Leu Leu Glu Pro Gly Val Pro Ala Gly Thr Cys Pro Lys Asp			
385	390	395	400
Tyr Val Glu Ile Asn Gly Glu Lys Tyr Cys Gly Glu Arg Ser Gln Phe			
405	410	415	
Val Val Thr Ser Asn Ser Asn Ile Thr Val Arg Phe His Ser Asp			
420	425	430	
Gln Ser Tyr Thr Asp Thr Gly Phe Leu Ala Glu Tyr Leu Ser Tyr Asp			
435	440	445	
Ser Ser Asp Pro Cys Pro Gly Gln Phe Thr Cys Arg Thr Gly Arg Cys			
450	455	460	
Ile Arg Lys Glu Leu Arg Cys Asp Gly Trp Ala Asp Cys Thr Asp His			
465	470	475	480
Ser Asp Glu Leu Asn Cys Ser Cys Asp Ala Gly His Gln Phe Thr Cys			
485	490	495	
Lys Asn Lys Phe Cys Lys Pro Leu Phe Trp Val Cys Asp Ser Val Asn			
500	505	510	
Asp Cys Gly Asp Asn Ser Asp Glu Gln Gly Cys Ser Cys Pro Ala Gln			
515	520	525	
Thr Phe Arg Cys Ser Asn Gly Lys Cys Leu Ser Lys Ser Gln Gln Cys			
530	535	540	
Asn Gly Lys Asp Asp Cys Gly Asp Gly Ser Asp Glu Ala Ser Cys Pro			
545	550	555	560
Lys Val Asn Val Val Thr Cys Thr Lys His Thr Tyr Arg Cys Leu Asn			
565	570	575	
Gly Leu Cys Leu Ser Lys Gly Asn Pro Glu Cys Asp Gly Lys Glu Asp			
580	585	590	
Cys Ser Asp Gly Ser Asp Glu Lys Asp Cys Asp Cys Gly Leu Arg Ser			
595	600	605	
Phe Thr Arg Gln Ala Arg Val Val Gly Gly Thr Asp Ala Asp Glu Gly			
610	615	620	
Glu Trp Pro Trp Gln Val Ser Leu His Ala Leu Gly Gln Gly His Ile			
625	630	635	640
Cys Gly Ala Ser Leu Ile Ser Pro Asn Trp Leu Val Ser Ala Ala His			
645	650	655	
Cys Tyr Ile Asp Asp Arg Gly Phe Arg Tyr Ser Asp Pro Thr Gln Trp			
660	665	670	
Thr Ala Phe Leu Gly Leu His Asp Gln Ser Gln Arg Ser Ala Pro Gly			
675	680	685	
Val Gln Glu Arg Arg Leu Lys Arg Ile Ile Ser His Pro Phe Phe Asn			
690	695	700	
Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu Leu Glu Leu Glu Lys Pro			
705	710	715	720
Ala Glu Tyr Ser Ser Met Val Arg Pro Ile Cys Leu Pro Asp Ala Ser			

725	730	735	
His Val Phe Pro Ala Gly Lys Ala Ile Trp Val Thr Gly Trp Gly His			
740	745	750	
Thr Gln Tyr Gly Gly Thr Gly Ala Leu Ile Leu Gln Lys Gly Glu Ile			
755	760	765	
Arg Val Ile Asn Gln Thr Thr Cys Glu Asn Leu Leu Pro Gln Gln Ile			
770	775	780	
Thr Pro Arg Met Met Cys Val Gly Phe Leu Ser Gly Gly Val Asp Ser			
785	790	795	800
Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser Ser Val Glu Ala Asp Gly			
805	810	815	
Arg Ile Phe Gln Ala Gly Val Val Ser Trp Gly Asp Gly Cys Ala Gln			
820	825	830	
Arg Asn Lys Pro Gly Val Tyr Arg Leu Pro Leu Phe Arg Asp Trp			
835	840	845	
Ile Lys Glu Asn Thr Gly Val			
850	855		

<210> 3
<211> 3147
<212> DNA
<213> Homo Sapien

<220>
<221> CDS
<222> (1865) ... (2590)
<223> Nucleic acid sequence of protease domain of MTSP1

<400> 3
tcaagagcgg cctcggggta ccatggggag cgatcgccc cgcaaggccg gagggggccc 60
gaaggacttc ggcgcgggac tcaagtacaa ctccccggcac gagaaaagtga atggcttgaa 120
ggaaggcgtg gagttctcgc cagtcaccaa cgtcaagaag gtggaaaagc atggcccgaa 180
gcgctgggtg gtgctggcag ccgtgctgtat cggccttc ttggcttgc tgggatcg 240
cttcctgggtg tggcatttgc agtaccggaa cgtgcgtgtc cagaaggctt tcaatggcta 300
catgaggatc acaaatttggaa atttttgttgc tgcctacgg aactccaact ccacttagtt 360
tgtaaggcctg gccagcaagg tgaaggacgc gctgaagctg ctgtacagcg gagtcccatt 420
cctggggccc taccacaagg agtcggctgt gacggccttc agcgaggccg gcgcatcg 480
ctactactgg tctgagttca gcatcccgca gcacctggtg gaggaggccg agcgcgtcat 540
ggccgaggag cgcgtagttca tgctggccccc gcgggcgcgc tccctgaagt cctttgtgg 600
cacctcgtg gtggcttcc ccacggactc caaaacagta cagaggaccc aggacaacag 660
ctgcagctt ggcctgcacg ccccggtgt ggagctgtat cgcttcacca cgcccggtt 720
ccctgacagc ccctaccccg ctcatgccccg ctgcctgttgc gcccctgcggg gggacgccc 780
ctcagtgtcg agcctcacct tccgcagctt tgaccttgcg tcctgcacg agcgcggcag 840
cgacctgggtg acgggttaca acaccctgag ccccatggag ccccaacgccc tggtgca 900
gtgtggacc taccctccct cctacaacccat gacctccac tcctcccaga acgttctgt 960
catcacactg ataaccaaca ctgagcggcg gcatccggc tttgaggcca cttttttcca 1020
gctgcctagg atgagcgtgttgcggcccttacgtaaa gcccaggggaa cattcaacag 1080
ccctactac ccaggccact acccaccacaa cattgactgc acatggaaaca ttgaggtgcc 1140
caacaaccag catgttggggacttcaaa atttttttgc ctgctggagc ccggcgtgcc 1200
tgccggacc tggcccaagg actacgttgc gatcaatggg gagaataact gcccggag 1260
gtcccgatcc gtcgtcacca gcaacacgcaa caagatcaca gttcgcttcc actcagatca 1320
gtctacacc gacaccggct tcttagctga atacccctcc tacgactcca gtgacccatg 1380
cccgccggcag ttcacgtgcc gcacggggccg gtgtatccgg aaggagctgc gctgtgtatgg 1440
ctggccgac tgcacccgacc acagcgatgt gctcaactgc agttgcgacg ccggccacca 1500
gttcacgtgc aagaacaagt totgcaagcc cctcttctgg gtctgcgaca gtgtgaacga 1560
ctgcggagac aacagcgacg agcagggtgt cagttgtccg gcccagaccc tcaggtgttc 1620
caatgggaag tgcctctcga aaagccagca gtgcaatggg aaggacgtgtgggacgg 1680
gtccgacgag gcctctgccc ccaagggtgaa cgtcgtcaact ttttacaaaac acacccatcg 1740
ctgcctcaat gggctctgttgc tgagcaaggg caaccctgag ttttacaaaac acacccatcg 1800
tagcgacggc tcagatgaga aggactgcga ctgtggctg cggtcattca cgagacaggc 1860

tcgt gtt gtt qgg ggc acg gat gcg gat gag ggc gag tgg ccc tgg cag Val Val Gly Gly Thr Asp Ala Asp Glu Gly Glu Trp Pro Trp Gln 1 5 10 15	1909
gta agc ctg cat gct ctg ggc cag ggc cac atc tgc ggt gct tcc ctc Val Ser Leu His Ala Leu Gly Gln Gly His Ile Cys Gly Ala Ser Leu 20 25 30	1957
atc tct ccc aac tgg ctg gtc tct gcc gca cac tgc tac atc gat gac Ile Ser Pro Asn Trp Leu Val Ser Ala Ala His Cys Tyr Ile Asp Asp 35 40 45	2005
aga gga ttc agg tac tca gac ccc acg cag tgg acg gcc ttc ctg ggc Arg Gly Phe Arg Tyr Ser Asp Pro Thr Gln Trp Thr Ala Phe Leu Gly 50 55 60	2053
ttg cac gac cag agc cag cgc agc gcc cct ggg gtg cag gag cgc agg Leu His Asp Gln Ser Gln Arg Ser Ala Pro Gly Val Gln Glu Arg Arg 65 70 75	2101
ctc aag cgc atc atc tcc cac ccc ttc ttc aat gac ttc acc ttc gac Leu Lys Arg Ile Ile Ser His Pro Phe Phe Asn Asp Phe Thr Phe Asp 80 85 90 95	2149
tat gac atc gcg ctg ctg gag ctg gag aaa ccg gca gag tac agc tcc Tyr Asp Ile Ala Leu Leu Glu Leu Lys Pro Ala Glu Tyr Ser Ser 100 105 110	2197
atg gtg cgg ccc atc tgc ctg ccg gac gcc tcc cat gtc ttc cct gcc Met Val Arg Pro Ile Cys Leu Pro Asp Ala Ser His Val Phe Pro Ala 115 120 125	2245
ggc aag gcc atc tgg gtc acg ggc tgg gga cac acc cag tat gga ggc Gly Lys Ala Ile Trp Val Thr Gly Trp Gly His Thr Gln Tyr Gly Gly 130 135 140	2293
act ggc gcg ctg atc ctg caa aag ggt gag atc cgc gtc atc aac cag Thr Gly Ala Leu Ile Leu Gln Lys Gly Glu Ile Arg Val Ile Asn Gln 145 150 155	2341
acc acc tgc gag aac ctc ctg ccg cag cag atc acg ccg cgc atg atg Thr Thr Cys Glu Asn Leu Leu Pro Gln Gln Ile Thr Pro Arg Met Met 160 165 170 175	2389
tgc gtg ggc ttc ctc agc ggc ggc gtg gac tcc tgc cag ggt gat tcc Cys Val Gly Phe Leu Ser Gly Gly Val Asp Ser Cys Gln Gly Asp Ser 180 185 190	2437
ggg gga ccc ctg tcc agc gtg gag ggc gat ggg cgg atc ttc cag gcc Gly Gly Pro Leu Ser Ser Val Glu Ala Asp Gly Arg Ile Phe Gln Ala 195 200 205	2485
ggt gtg gtg agc tgg gga gac ggc tgc gct cag agg aac aag cca ggc Gly Val Val Ser Trp Gly Asp Gly Cys Ala Gln Arg Asn Lys Pro Gly 210 215 220	2533
gtg tac aca agg ctc cct ctg ttt cgg gac tgg atc aaa gag aac act Val Tyr Thr Arg Leu Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr 225 230 235	2581

ggg gta tag gggccggggc cacccaaatg tgtacacctg cggggccacc 2630
Gly Val *
240

catcgccccac cccagtgtgc acgcctgcag gctggagact ggaccgctga ctgcaccagc 2690
gccccccagaa catacaactgt gaactcaatc tccagggtctc caaatctgcc tagaaaaacct 2750
ctcgcccttccct cagcctccaa agtggagctg ggaggttagaa ggggaggaca ctgggtggttc 2810
tactgaccca actgggggca aagggttgaa gacacagcct ccccccgcag ccccaagctg 2870
ggccgaggcg cggtttgtgt tatctgcctc ccctgtctgt aaggagcagc gggAACGGAG 2930
cttcggagcc tcctcagtgta aggtgggtgg gctgcccggat ctgggctgtg gggcccttgg 2990
gccacgcctct tgaggaagcc caggctcgga ggaccctgga aaacagacgg gtctgagact 3050
gaaattgtttt accagctcc cagggtggac ttcaagtgtgt gtatttgtt aaatggtaa 3110
aacaatttat ttcttttaa aaaaaaaaaaaa aaaaaaaaaa 3147

<210> 4
<211> 241
<212> PRT
<213> Homo Sapien

<400> 4
Val Val Gly Gly Thr Asp Ala Asp Glu Gly Glu Trp Pro Trp Gln Val 15
1 5 10 15
Ser Leu His Ala Leu Gly Gln Gly His Ile Cys Gly Ala Ser Leu Ile 20 25 30
20 25 30
Ser Pro Asn Trp Leu Val Ser Ala Ala His Cys Tyr Ile Asp Asp Arg 35 40 45
35 40 45
Gly Phe Arg Tyr Ser Asp Pro Thr Gln Trp Thr Ala Phe Leu Gly Leu 50 55 60
50 55 60
His Asp Gln Ser Gln Arg Ser Ala Pro Gly Val Gln Glu Arg Arg Leu 65 70 75 80
65 70 75 80
Lys Arg Ile Ile Ser His Pro Phe Phe Asn Asp Phe Thr Phe Asp Tyr 85 90 95
85 90 95
Asp Ile Ala Leu Leu Glu Leu Glu Lys Pro Ala Glu Tyr Ser Ser Met 100 105 110
100 105 110
Val Arg Pro Ile Cys Leu Pro Asp Ala Ser His Val Phe Pro Ala Gly 115 120 125
115 120 125
Lys Ala Ile Trp Val Thr Gly Trp Gly His Thr Gln Tyr Gly Thr 130 135 140
130 135 140
Gly Ala Leu Ile Leu Gln Lys Gly Glu Ile Arg Val Ile Asn Gln Thr 145 150 155 160
145 150 155 160
Thr Cys Glu Asn Leu Leu Pro Gln Gln Ile Thr Pro Arg Met Met Cys 165 170 175
165 170 175
Val Gly Phe Leu Ser Gly Gly Val Asp Ser Cys Gln Gly Asp Ser Gly 180 185 190
180 185 190
Gly Pro Leu Ser Ser Val Glu Ala Asp Gly Arg Ile Phe Gln Ala Gly 195 200 205
195 200 205
Val Val Ser Trp Gly Asp Gly Cys Ala Gln Arg Asn Lys Pro Gly Val 210 215 220
210 215 220
Tyr Thr Arg Leu Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr Gly 225 230 235 240
225 230 235 240
Val

<210> 5
<211> 2293
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)...(2259)

<223> CVSP16 Full Length cDNA

<400> 5 atg gcc cgg cag ctg ctc ccc ctt gtg gtg ctt gtc atc agt ccc Met Ala Arg Gln Leu Leu Leu Pro Leu Val Val Leu Val Ile Ser Pro 1 5 10 15	48
atc cca gga gcc ttc cag gac tca gct ctc agt cct acc cag gaa gaa Ile Pro Gly Ala Phe Gln Asp Ser Ala Leu Ser Pro Thr Gln Glu Glu 20 25 30	96
cct gaa gat ctg gac tgc ggg cgc cct gag ccc tcg gcc cgc atc gtg Pro Glu Asp Leu Asp Cys Gly Arg Pro Glu Pro Ser Ala Arg Ile Val 35 40 45	144
ggg ggc tca aac gcg cag ccg ggc acc tgg cct tgg caa gtg agc ctg Gly Gly Ser Asn Ala Gln Pro Gly Thr Trp Pro Trp Gln Val Ser Leu 50 55 60	192
cac cat gga ggt ggc cac atc tgc ggg ggc tcc ctc atc gcc ccc tcc His His Gly Gly His Ile Cys Gly Ser Leu Ile Ala Pro Ser 65 70 75 80	240
tgg gtc ctc tcc gcc gct cac tgt ttc atg acg aat ggg acg ctg gag Trp Val Leu Ser Ala Ala His Cys Phe Met Thr Asn Gly Thr Leu Glu 85 90 95	288
ccc gcg gcc gag tgg tcg gta ctg ctg ggc gtg cac tcc cag gac ggg Pro Ala Ala Glu Trp Ser Val Leu Leu Gly Val His Ser Gln Asp Gly 100 105 110	336
ccc ctg gac ggc gcg cac acc cgc gca gtg gcc gcc atc gtg gtg ccg Pro Leu Asp Gly Ala His Thr Arg Ala Val Ala Ala Ile Val Val Pro 115 120 125	384
gcc aac tac agc caa gtg gag ctg ggc gac ctg gcc ctg ctg cgc Ala Asn Tyr Ser Gln Val Glu Leu Gly Ala Asp Leu Ala Leu Leu Arg 130 135 140	432
ctg gcc tca ccc gcc agc ctg ggc ccc gcc gtg tgg cct gtc tgc ctg Leu Ala Ser Pro Ala Ser Leu Gly Pro Ala Val Trp Pro Val Cys Leu 145 150 155 160	480
ccc cgc gcc tca cac cgc ttc gtg cac ggc acc gcc tgc tgg gcc acc Pro Arg Ala Ser His Arg Phe Val His Gly Thr Ala Cys Trp Ala Thr 165 170 175	528
ggc tgg gga gac gtc cag gag gca gat cct ctg cct ctc ccc tgg gtg Gly Trp Gly Asp Val Gln Glu Ala Asp Pro Leu Pro Leu Pro Trp Val 180 185 190	576
cta cag gaa gtg gag cta agg ctg ctg ggc gag gcc acc tgt caa tgt Leu Gln Glu Val Glu Leu Arg Leu Leu Gly Glu Ala Thr Cys Gln Cys 195 200 205	624
ctc tac agc cag ccc ggt ccc ttc aac ctc act ctc cag ata ttg cca Leu Tyr Ser Gln Pro Gly Pro Phe Asn Leu Thr Leu Gln Ile Leu Pro 210 215 220	672
ggg atg ctg tgt gct ggc tac cca ggg ggc cgc agg gac acc tgc cag	720

Gly	Met	Leu	Cys	Ala	Gly	Tyr	Pro	Gly	Gly	Arg	Arg	Asp	Thr	Cys	Gln		
225				230				235				240					
ggt	gac	tct	ggg	ggg	ccc	ctg	gtc	tgt	gag	gaa	ggc	ggc	cgc	tgg	ttc	768	
Gly	Asp	Ser	Gly	Gly	Pro	Leu	Val	Cys	Glu	Glu	Gly	Gly	Arg	Trp	Phe		
245					250				255								
cag	gca	gga	atc	acc	agc	ttt	ggc	ttt	ggc	tgt	gga	cg	aga	aac	cgc	816	
Gln	Ala	Gly	Ile	Thr	Ser	Phe	Gly	Phe	Gly	Cys	Gly	Arg	Arg	Asn	Arg		
260					265				270								
cct	gga	gtt	ttc	act	gct	gtg	gct	acc	tat	gag	gca	tgg	ata	cg	gag	864	
Pro	Gly	Val	Phe	Thr	Ala	Val	Ala	Thr	Tyr	Glu	Ala	Trp	Ile	Arg	Glu		
275					280				285								
cag	gtg	atg	gg	tca	gag	cct	ggg	cct	gcc	ttt	ccc	acc	cag	ccc	cag	912	
Gln	Val	Met	Gly	Ser	Glu	Pro	Gly	Pro	Ala	Phe	Pro	Thr	Gln	Pro	Gln		
290					295				300								
aag	acc	cag	tca	gat	ccc	cag	gag	ccc	agg	gag	gag	aac	tgc	acc	att	960	
Lys	Thr	Gln	Ser	Asp	Pro	Gln	Glu	Pro	Arg	Glu	Glu	Asn	Cys	Thr	Ile		
305					310				315			320					
gcc	ctg	cct	gag	tgc	ggg	aag	gcc	ccg	cg	cca	ggg	gcc	tgg	ccc	tgg	1008	
Ala	Leu	Pro	Glu	Cys	Gly	Lys	Ala	Pro	Arg	Pro	Gly	Ala	Trp	Pro	Trp		
325					330				335								
gag	gcc	cag	gtg	atg	gtg	cca	gga	tcc	aga	ccc	tgc	cat	ggg	g	ctg	1056	
Glu	Ala	Gln	Val	Met	Val	Pro	Gly	Ser	Arg	Pro	Cys	His	Gly	Ala	Leu		
340					345				350								
gtg	tct	gaa	agc	tgg	gtc	ttg	gca	cct	gcc	agc	tgc	ttt	ctg	gac	ccg	1104	
Val	Ser	Glu	Ser	Trp	Val	Leu	Ala	Pro	Ala	Ser	Cys	Phe	Leu	Asp	Pro		
355					360				365								
aac	agc	tcc	gac	agc	cca	ccc	cgc	gac	ctc	gac	gcc	tgg	cgc	gtg	ctg	1152	
Asn	Ser	Ser	Asp	Ser	Pro	Pro	Arg	Asp	Leu	Asp	Ala	Trp	Arg	Val	Leu		
370					375				380								
ctg	ccc	tgc	cac	ccg	cgc	gag	cg	gtg	g	cg	cg	ctg	gtg	cag	cac	1200	
Leu	Pro	Ser	His	Pro	Arg	Ala	Glu	Arg	Val	Ala	Arg	Leu	Val	Gln	His		
385					390				395			400					
gag	aac	gct	tcg	tgg	gac	aa	cc	cc	gac	ctg	g	cg	ctg	cag	ctg	1248	
Glu	Asn	Ala	Ser	Trp	Asp	Asn	Ala	Pro	Asp	Leu	Ala	Leu	Leu	Gln	Leu		
405					410				415								
cg	ac	cc	gt	aa	ct	ag	tg	ct	cg	cc	gt	tg	ct	ca	cc	1296	
Arg	Thr	Pro	Val	Asn	Leu	Ser	Ala	Ala	Ser	Arg	Pro	Val	Cys	Leu	Pro		
420					425				430								
cac	ccg	gaa	cac	tac	ttc	ct	cc	gg	ag	cg	tg	cg	ct	gg	cc	1344	
His	Pro	Glu	His	Tyr	Phe	Leu	Pro	Gly	Ser	Arg	Cys	Arg	Leu	Ala	Arg		
435					440				445								
tgg	ggc	cg	gg	gaa	cc	gc	ctt	ggc	cc	gg	cg	ct	tg	gag	cg	1392	
Trp	Gly	Arg	Gly	Gly	Gl	Pro	Ala	Leu	Gly	Pro	Gly	Ala	Leu	Leu	Glu	Ala	
450					455				460								
gag	ctg	tta	ggc	ggc	tgg	tgg	tgc	cac	tgc	ctg	tac	ggc	cgc	cag	gg	1440	

Glu	Leu	Leu	Gly	Gly	Trp	Trp	Cys	His	Cys	Leu	Tyr	Gly	Arg	Gln	Gly		
465					470				475				480				
gcg	gca	gta	ccg	ctg	ccc	gga	gac	ccg	cac	gcg	ctc	tgc	cct	gcc		1488	
Ala	Ala	Val	Pro	Leu	Pro	Gly	Asp	Pro	Pro	His	Ala	Leu	Cys	Pro	Ala		
														495			
485									490								
tac	cag	gaa	aag	gag	gag	gtg	ggc	agc	tgc	tgg	aat	gac	tcg	cgt	tgg		1536
Tyr	Gln	Glu	Lys	Glu	Val	Gly	Ser	Cys	Trp	Asn	Asp	Ser	Arg	Trp			
														510			
500									505								
agc	ctt	ttg	tgc	cag	gag	gag	ggg	acc	tgg	ttt	ctg	gct	gga	atc	aga		1584
Ser	Leu	Leu	Cys	Gln	Glu	Glu	Gly	Thr	Trp	Phe	Leu	Ala	Gly	Ile	Arg		
														525			
515									520								
gac	ttt	ccc	agt	ggc	tgt	cta	cgt	ccc	cga	gcc	ttc	ttc	cct	ctg	cag		1632
Asp	Phe	Pro	Ser	Gly	Cys	Leu	Arg	Pro	Arg	Ala	Phe	Phe	Pro	Leu	Gln		
														540			
530								535									
act	cat	ggc	cca	tgg	atc	agc	cat	gtg	act	cgg	gga	gcc	tac	ctg	gag		1680
Thr	His	Gly	Pro	Trp	Ile	Ser	His	Val	Thr	Arg	Gly	Ala	Tyr	Leu	Glu		
														560			
545								550									
gac	cag	cta	gcc	tgg	gac	tgg	ggc	cct	gat	ggg	gag	gag	act	gag	aca		1728
Asp	Gln	Leu	Ala	Trp	Asp	Trp	Gly	Pro	Asp	Gly	Glu	Glu	Thr	Glu	Thr		
														575			
565									570								
cag	act	tgt	ccc	cca	cac	aca	gag	cat	ggt	gcc	tgt	ggc	ctg	cg	ctg		1776
Gln	Thr	Cys	Pro	Pro	His	Thr	Glu	His	Gly	Ala	Cys	Gly	Leu	Arg	Leu		
														590			
580									585								
gag	gct	gct	cca	gtg	ggg	gtc	ctg	tgg	ccc	tgg	ctg	gca	gag	gtg	cat		1824
Glu	Ala	Ala	Pro	Val	Gly	Val	Leu	Trp	Pro	Trp	Leu	Ala	Glu	Val	His		
														605			
595								600									
gtg	gct	ggt	gat	cga	gtc	tgc	act	ggg	atc	ctc	ctg	gcc	cca	ggc	tgg		1872
Val	Ala	Gly	Asp	Arg	Val	Cys	Thr	Gly	Ile	Leu	Leu	Ala	Pro	Gly	Trp		
														620			
610									615								
gtc	ctg	gca	gcc	act	cac	tgt	gtc	ctc	agg	cca	ggc	tct	aca	aca	gtg		1920
Val	Leu	Ala	Ala	Thr	His	Cys	Val	Leu	Arg	Pro	Gly	Ser	Thr	Thr	Val		
														640			
625									630								
cct	tac	att	gaa	gtg	tat	ctg	ggc	gca	ggg	gcc	agc	tcc	ctc	cca		1968	
Pro	Tyr	Ile	Glu	Val	Tyr	Leu	Gly	Arg	Ala	Gly	Ala	Ser	Ser	Leu	Pro		
														655			
645									650								
cag	ggc	cac	cag	atg	acc	tca	gca	ccg	ccc	ctc	ctg	tgc	cag	atg	acg		2016
Gln	Gly	His	Gln	Met	Thr	Ser	Ala	Pro	Pro	Leu	Leu	Cys	Gln	Met	Thr		
														670			
660									665								
gaa	ggg	tcc	tgg	atc	ctc	gtg	ggc	atg	gct	caa	ggg	agc	cg	gag		2064	
Glu	Gly	Ser	Trp	Ile	Leu	Val	Gly	Met	Ala	Val	Gln	Gly	Ser	Arg	Glu		
														685			
675									680								
ctg	ttt	gct	gcc	att	ggt	cct	gaa	gag	gcc	tgg	atc	tcc	cag	aca	gtg		2112
Leu	Phe	Ala	Ala	Ile	Gly	Pro	Glu	Glu	Ala	Trp	Ile	Ser	Gln	Thr	Val		
														700			
690									695								
gga	gag	gcc	aac	ttc	ctg	ccc	ccc	agt	ggc	tcc	cca	cac	tgg	ccc	act		2160

Gly Glu Ala Asn Phe Leu Pro Pro Ser Gly Ser Pro His Trp Pro Thr			
705	710	715	720
gga ggc agc aat ctc tgc ccc cca gaa ctg gcc aag gcc tcg gga tcc			2208
Gly Gly Ser Asn Leu Cys Pro Pro Glu Leu Ala Lys Ala Ser Gly Ser			
725	730	735	
ccg cat gca gtc tac ttc ctg ctc ctg act ctc ctg atc cag agc			2256
Pro His Ala Val Tyr Phe Leu Leu Leu Thr Leu Leu Ile Gln Ser			
740	745	750	
tga gggcttaggg tcccaagcacc acttccccct tctc			2293
*			

<210> 6
<211> 752
<212> PRT
<213> Artificial Sequence

<220>
<223> CVSP16 Full Length Protein

<400> 6			
Met Ala Arg Gln Leu Leu Leu Pro Leu Val Val Leu Val Ile Ser Pro			
1	5	10	15
Ile Pro Gly Ala Phe Gln Asp Ser Ala Leu Ser Pro Thr Gln Glu Glu			
20	25	30	
Pro Glu Asp Leu Asp Cys Gly Arg Pro Glu Pro Ser Ala Arg Ile Val			
35	40	45	
Gly Gly Ser Asn Ala Gln Pro Gly Thr Trp Pro Trp Gln Val Ser Leu			
50	55	60	
His His Gly Gly His Ile Cys Gly Ser Leu Ile Ala Pro Ser			
65	70	75	80
Trp Val Leu Ser Ala Ala His Cys Phe Met Thr Asn Gly Thr Leu Glu			
85	90	95	
Pro Ala Ala Glu Trp Ser Val Leu Leu Gly Val His Ser Gln Asp Gly			
100	105	110	
Pro Leu Asp Gly Ala His Thr Arg Ala Val Ala Ala Ile Val Val Pro			
115	120	125	
Ala Asn Tyr Ser Gln Val Glu Leu Gly Ala Asp Leu Ala Leu Leu Arg			
130	135	140	
Leu Ala Ser Pro Ala Ser Leu Gly Pro Ala Val Trp Pro Val Cys Leu			
145	150	155	160
Pro Arg Ala Ser His Arg Phe Val His Gly Thr Ala Cys Trp Ala Thr			
165	170	175	
Gly Trp Gly Asp Val Gln Glu Ala Asp Pro Leu Pro Leu Pro Trp Val			
180	185	190	
Leu Gln Glu Val Glu Leu Arg Leu Leu Gly Glu Ala Thr Cys Gln Cys			
195	200	205	
Leu Tyr Ser Gln Pro Gly Pro Phe Asn Leu Thr Leu Gln Ile Leu Pro			
210	215	220	
Gly Met Leu Cys Ala Gly Tyr Pro Gly Gly Arg Arg Asp Thr Cys Gln			
225	230	235	240
Gly Asp Ser Gly Gly Pro Leu Val Cys Glu Glu Gly Arg Trp Phe			
245	250	255	
Gln Ala Gly Ile Thr Ser Phe Gly Phe Gly Cys Gly Arg Arg Asn Arg			
260	265	270	
Pro Gly Val Phe Thr Ala Val Ala Thr Tyr Glu Ala Trp Ile Arg Glu			

275	280	285
Gln Val Met Gly Ser Glu Pro	Gly Pro Ala Phe Pro	Thr Gln Pro Gln
290	295	300
Lys Thr Gln Ser Asp Pro	Gln Glu Pro Arg	Glu Asn Cys Thr Ile
305	310	315
Ala Leu Pro Glu Cys Gly Lys	Ala Pro Arg Pro	Gly Ala Trp Pro Trp
325	330	335
Glu Ala Gln Val Met Val Pro	Gly Ser Arg Pro	Cys His Gly Ala Leu
340	345	350
Val Ser Glu Ser Trp Val	Leu Ala Pro Ala Ser	Cys Phe Leu Asp Pro
355	360	365
Asn Ser Ser Asp Ser Pro	Pro Arg Asp	Leu Asp Ala Trp Arg Val Leu
370	375	380
Leu Pro Ser His Pro Arg	Ala Glu Arg Val	Ala Arg Leu Val Gln His
385	390	395
Glu Asn Ala Ser Trp Asp	Asn Ala Pro Asp	Leu Ala Leu Leu Gln Leu
405	410	415
Arg Thr Pro Val Asn Leu Ser	Ala Ala Ser Arg	Pro Val Cys Leu Pro
420	425	430
His Pro Glu His Tyr Phe	Leu Pro Gly Ser Arg	Cys Arg Leu Ala Arg
435	440	445
Trp Gly Arg Gly Glu Pro	Ala Leu Gly Pro Gly	Ala Leu Leu Glu Ala
450	455	460
Glu Leu Leu Gly Gly Trp	Trp Cys His Cys	Leu Tyr Gly Arg Gln Gly
465	470	475
Ala Ala Val Pro Leu Pro	Gly Asp Pro Pro	His Ala Leu Cys Pro Ala
485	490	495
Tyr Gln Glu Lys Glu Glu	Val Gly Ser Cys	Trp Asn Asp Ser Arg Trp
500	505	510
Ser Leu Leu Cys Gln Glu	Gly Thr Trp Phe	Leu Ala Gly Ile Arg
515	520	525
Asp Phe Pro Ser Gly Cys	Leu Arg Pro Arg	Ala Phe Phe Pro Leu Gln
530	535	540
Thr His Gly Pro Trp Ile	Ser His Val Thr	Arg Gly Ala Tyr Leu Glu
545	550	555
Asp Gln Leu Ala Trp Asp	Trp Gly Pro Asp	Gly Glu Glu Thr Glu Thr
565	570	575
Gln Thr Cys Pro Pro His	Thr Glu His Gly	Ala Cys Gly Leu Arg Leu
580	585	590
Glu Ala Ala Pro Val Gly	Val Leu Trp Pro	Trp Leu Ala Glu Val His
595	600	605
Val Ala Gly Asp Arg Val	Cys Thr Gly Ile	Leu Leu Ala Pro Gly Trp
610	615	620
Val Leu Ala Ala Thr His	Cys Val Leu Arg	Pro Gly Ser Thr Thr Val
625	630	635
Pro Tyr Ile Glu Val Tyr	Leu Gly Arg Ala	Gly Ala Ser Ser Leu Pro
645	650	655
Gln Gly His Gln Met Thr	Ser Ala Pro Pro	Leu Leu Cys Gln Met Thr
660	665	670
Glu Gly Ser Trp Ile Leu	Val Gly Met Ala Val	Gln Gly Ser Arg Glu
675	680	685
Leu Phe Ala Ala Ile Gly	Pro Glu Glu Ala Trp	Ile Ser Gln Thr Val
690	695	700
Gly Glu Ala Asn Phe	Leu Pro Pro Ser	Gly Ser Pro His Trp Pro Thr
705	710	715
Gly Gly Ser Asn Leu Cys	Pro Pro Glu Leu	Ala Lys Ala Ser Gly Ser
725	730	735
Pro His Ala Val Tyr Phe	Leu Leu Leu	Leu Thr Leu Ile Gln Ser
740	745	750

```

<210> 7
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 7
ccctctgggt agccagcaca cagcatc 27

<210> 8
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 8
gccatcggt tgccggccaa ctacag 26

<210> 9
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 9
gcacacagca tccctggcaa tatctgg 27

<210> 10
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 10
cggccaacta cagccaagtg gagctg 26

<210> 11
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 11
atggccccggc agctgctcct ccccctttgt 30

<210> 12
<211> 28
<212> DNA
<213> Artificial Sequence

```

<220>
<223> Primer

<400> 12
cggtccgg gcaggaagta gtgttccg

28

<210> 13
<211> 9276
<212> DNA
<213> Pichi

<400> 13
agatctaaca tccaaagacg aaagggttcaa taaaacctt ttgcacatcg acatccacag 60
gtccattctc acacataatg gccaaacgcg acaggagggg atacactagc agcagaccgt 120
tgcaaacgca ggacctccac tcctttctc ctcacacccc actttgcca tcgaaaaacc 180
agcccaagta tttgggttga ttggagctcg ctcattccaa ttccatctat taggctacta 240
acacatgcac ttatttagcc tgtctatcct ggccccccctg gcgagggttca tggttgttta 300
tttccgaatg caacaagctc cgcatcacac ccgaacatca ctccagatga gggcttctg 360
agtgtgggt caaatagttt catgtcccc aatggccca aaactgacag tttaaacgct 420
gtcttggAAC ctaatatgac aaaagcgtga tctcatccaa gatgaactaa gtttgggtcg 480
ttgaaatgct aacggccagt tggtaaaaaa gaaacttcca aaagtgcgccaa taccgttgt 540
cttgggttgc attgattgac gaatgctaa aaataatctc attaatgctt agcgcagct 600
ctctatcgct tctgaacccc ggtgcacctg tgccgaaacg caaatgggaa aacaccgct 660
tttggatgaa ttatgcattt tctccacatt gtatgttcc aagattctgg tggaaatact 720
gctgatagcc taacgttcat gatcaaaaatt taactgttcc aacccttact tgacagcaat 780
atataaacag aaggaaagctg ccctgtctta aacctttttt ttatcatca ttattagctt 840
actttatcaa ttgcgactgg ttccaaatgaa caaactaattt attcgaagga tccaaacgat 900
caacttgaga agatcaaaaaa acaactaattt attcgaagga tccaaacgat gagatttcct 960
tcaattttt ctgcagttt attcgcagca tcctccgcatt tagtgcgtcc agtcaacact 1020
acaacaaagaa atgaaaacggc acaaattccg gctgaagctg tcatcggtta ctcagattta 1080
gaaggggatt tcgatgttgc tggtttgcca ttttccaaaca gcacaaataaa cgggttattg 1140
tttataaataa ctactattgc cagcattgct gctaaagaag aagggttatac tctcgagaaa 1200
agagaggctg aagcttacgt agaattccct agggcggccg cgaattaattt cgccttagac 1260
atgactgttc ctcaagttcaa gttgggact tacgagaaga ccgtcttgc tagattctaa 1320
tcaagaggat gtcagaatgc catttgcctg agagatgcag gcttcatttt tgatacttt 1380
ttatttgtaa cctatataatg atagatttt ttttgcattt ttgttcttc tcgtacgac 1440
ttgctcttga tcagcctatc tgcagctga tgaatatctt gtgttagggg tttggggaaa 1500
tcattcgagt ttgatgttt tctttgtatt tccacttctt cttcagagta cagaagatta 1560
agtgagaagt tcgtttgtgc aagcttacg ataagcttta atgcggtagt ttatcacagt 1620
taaattgtcta acgcagtcag gcaccgtgtc tggagccac tatcgtcatcc 1680
tcggcaccgt caccctggat gctgttaggc taggcttgggt tatgcccgtt ctggggggcc 1740
tcttgcggga tategtccat tccgacagca tcgccagtca ctatggcgtg ctgctagcgc 1800
tatatgcgtt gatgcaattt ctatgcgcac ccgttctcgg agcactgtcc gaccgcttgc 1860
gcccgcgccc agtcctgctc gcttcgtac ttggagccac tattcgtactac gcatgtcatgg 1920
cgaccacacc cgtccctgtgg atctatcgaa tctaaatgtt agttaaaatc tctaaataat 1980
taaataaagtcc ccaatgttccat catacgaacc ttaacagcat tgcggtgagc atctagaccc 2040
tcaacacgcg ccagatccat cactgcttgg ccaatgttccat tcaatgttcc tcaatgttcc 2100
tcttgcgaag tgatgaacctt ctggaaagggtt gcaatgttccat tcaatgttcc tcaatgttcc 2160
tatccgtacg ttggcaaaatg gtggttggta cgggaggagt aatctccaca tctctgttgc 2220
gagtaggcac caacaaacac agatccagcg tggatgttact gatcaacata agaagaagca 2280
ttctcgattt gcaggatcaa gtgttccagg tggatgttact gctgtactgt tgacatttc caaaggctgc 2340
tcgttagttt caaccgtat ggtttagag gcatcttcaa ttctggcaag aatcttgc 2400
cttggcaact gcacagctt gtttgcgtt gcatcttcaa ttctggcaag aatcttgc 2460
gtcatatcgaa cagccaaacac aatcacctgg gatcaacata catgttccat ttgagacaga 2520
aggttcgtgg ccaacaaatc tggatcagcg gatcaacata catgttccat ttgagacaga 2580
ggcccaagcg gcatgtcaat actacacagg gatcaacata catgttccat ttgagacaga 2640
gcagcgttaa cgaactgggt tcctggacca gatcaacata catgttccat ttgagacaga 2700
gttccgtaaag ccatacgac tactgcctgg gatcaacata catgttccat ttgagacaga 2760
ccaaccttgt gggcaacatg gatgacttct gatcaacata catgttccat ttgagacaga 2820
gatgaaaaaaaaa caatttctt gcaaccaggca gatcaacata catgttccat ttgagacaga 2880

gtggaggca gaattgcgg tccaccagga atatagaggc caacttctc aataggtctt 2940
 gaaaaacgag agcagactac accaggcca gtctcaactt gcaacgtctc cgtagttga 3000
 gttcatgga atttcctgac gttatctata gagagatcaa tgctctt aacgttatct 3060
 ggcaattgca taagttcc tcggaaagga gcttctaaca caggtgtctt caaagcgact 3120
 ccatcaaact tggcagttag ttctaaaagg gcttgac catttgacg aacattgtcg 3180
 acaatggtt tgactaattc cataatctgt tccgtttctt ggataggacg acgaaggca 3240
 tcttcattt cttgtgagga ggccttagaa acgtcaattt tgccacaattc aatacgacct 3300
 tcagaaggga ctcttttagg ttggattct tcttaggtt gttccttggt gtatccggc 3360
 ttggatctc ctcccttctt agtgacctt agggacttca tatccaggtt tctctccacc 3420
 tcgtccaaacg tcacaccgtt ctggcacat ctaactaatg caaaataaaa taagtca 3480
 catcccagg ctatatttc ctggattta gcttctgaa gttcatcagc ttccctcccta 3540
 attttagcgt tcaacaaaac ttctgtgtca aataaccgtt tggataaga accttctgga 3600
 gcattgtct tacgatccca caaggtggct tccatggctc taagaccctt tgattggca 3660
 aaacaggaag tgcgttccaa gtgacagaaa ccaacacctg ttgttcaac cacaatttc 3720
 aagcgtctc catcacaatc caattcgata cccagcaact tttaggtgc tccagatgt 3780
 gcacccctt accacaaacc gtgacgacga gattggtaga ctccagttt tgccttata 3840
 gcctccgaa tagactttt ggacgagttc accaggccc acgagtaat agaagagtca 3900
 gccacccaaag tagtgaatag accatgggg cggtcgttag tcaaagacgc caacaaaatt 3960
 tcactgacag ggaactttt gacatctca gaaagtctgtt attcgtatg caattgccga 4020
 gcatcaataa tggggattt accagaagca acatggaag tcacatctac caacttgcg 4080
 gtctcagaaa aagcataaaac agttctacta ccgcatttag tgaaactttt caaatcgccc 4140
 agtggagaag aaaaggccac agcgatacta gcattagcgg gcaaggatgc aacttata 4200
 accagggtcc tataaaaaac cctagcgctt gggatcatcc ttggacaac tcttctgccc 4260
 aaatcttagt cccaaatcac ttcatgtata ccattattgt acaactttag caagttgtcg 4320
 atcagctcct caaattgttc ctctgttaacg gatgactcaa ctgcacatt aacttga 4380
 tcagtcgatt gagtgaacct gatcagggtt tgagctgtt cagcagcata gggaaacacg 4440
 gctttctta cccaaactcaa ggaattatca aactctgca cacttgcgtt tgcaaggtagc 4500
 aaggaaatg tcataacttga agtggacag tgagtgttgtt ctggagaaat tctgaagccg 4560
 tattttattt atcagtgttgtt cagtcatacg gagatccctt acgcggacg catcggtggcc 4620
 gacctgcagg gggggggggg ggcgtgaggt ctgcctcggt aagaagggtt tgctgactca 4680
 taccaggcct gaatcgcccc atcatccagc cagaaagtga gggagccacg ttgtatgaga 4740
 gcttgggtt aggtggacca gttgggtattt ttggactttt gcttggccac ggaacggct 4800
 gcgttgcgg gaagatgcgt gatctgtatcc ttcaactcag caaaagttcg atttattcaa 4860
 caaagccgccc gtcccgtaa gtcagcgtaa tgctctgcca gtgttacaac caattaacca 4920
 attctgatta gaaaactca tcgagcatca aatggaaactg caatttattt atatcaggat 4980
 tataatacc atattttga aaaaggcggtt tctgtatga aggagaaaaac tcaccgaggc 5040
 agttccatag gatggcaaga tcctggatc ggtctgcgtt ccgcactctg ccaacatcaa 5100
 tacaacatattt taatcccccc tcgtcaaaaaa taagggttata aagtggaaaaa tcaccatgag 5160
 tgacgactga atccggtgag aatggaaaaa gctttagtgcatttccatc acttggatca 5220
 caggatattc ttctaatacc tggaaatgtt tttcccggtt gatcgacgtt gtgagtaacc 5280
 atgcatcatc aggatgtacgg ataaaatgtt tgatggcgtt aagaggcata aattccgtca 5340
 gtcgatcttgc tctgaccatc tcatctgtaa catcattggc aacgctactt tgccatgtt 5580
 tcagaaacaa ctctggcgca tcgggcttcc catacaatcg atagattgtc gcacctgatt 5640
 gcccgcatt atcgcgagcc catttatacc catataaattc agcatccatg ttggatattt 5700
 atcgccgcct cgagcaagac gttcccggtt gaatatggctt cataacaccc cttgtattac 5760
 tgtttatgtt agcagacatg ttattgttca atgatgtat atttttatct tgcgtatgt 5820
 aacatcgatg attttggatc acaacgtggc ttcccccggg cccccctgcag gtcggcatca 5880
 cggcgccac aggtgcgggtt gctggcgctt atatcgccga catcaccgtt gggaaagatc 5940
 gggctcgcca ttccgggtt atgagcgctt gttccgggtt gggatgggtt gcaggcccc 6000
 tggccggggg actgttgggc gcatcttccatc tgcgtgcacc attccctgcg gccggcggtgc 6060
 tcaacggcct caacctacta ctgggctgtt tcctaattgcg gggatgcgtt aagggagac 6120
 gtcgagtttccatc tatgatttgcg agtatggaa tggtgatacc cgcatttttc agtgcgttgc 6180
 ggttcctat cagattatgc ccaactaaag caaccggagg aggagattt atggtaaattt 6240
 tctctgactt ttggatccatca gtagactcga actgtgagac tatctcggtt atgacagcag 6300
 aaatgtccctt ctggagaca gtaaatgaag tcccaccaat aaagaaatcc ttgttattcag 6360
 gaacaaactt ctgtttcga actttttcg gtccttgcac tataaaatgtt agatggata 6420
 tgcgggttag gaatggagcg ggcaaatgtt taccttgcg accttcaaga ggtatgtttagg 6480

gttttagat actgatgcc aactcagtga caacgttgct attcggtca aaccattccg 6540
 aatccagaga aatcaaagg tttgtctac tattgatcca agcagtgcg gtctgaaac 6600
 tgacaatagt gtgctcggt tttgaggtca tctttgtatg aataaatcta gtcttgatc 6660
 taaaataatct tgacgagcca aggcgataaa tacccaaatc taaaactt taaaacgtt 6720
 aaaaggacaa gtatgtctgc ctgtttaaa ccccaaatca gctcgtagtc tgatccat 6780
 caacttgagg ggcactatc ttttttagag aaatttgcgg agatgcgata tcgagaaaaa 6840
 ggtacgctga ttttaaactg gaaattatc tcaagatc tgcctcgcc gtttcggta 6900
 tgacggtaa aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaa 6960
 ggtatggccgg agcagacaag cccgtcagg cgctcagcg ggtgttggcg ggtgtcggg 7020
 cgcagccatg acccagtca gtagcgatag cggagtgtat actggcttaa ctatgcggca 7080
 tcagagcaga ttgtactgag agtgcaccat atgcgggtgtg aaataccgca cagatgcgta 7140
 aggagaaaaat accgcatacg ggcgttcc cactcaaagg cgtaataacg gttatccaca 7260
 gtcgtccggc tgcggcgagc ggtatcagct tgagcaaaag gccagcaaaa ggccaggaac 7320
 gaatcagggg ataacgcagg aaagaacatg cataggctcc gccccctga cgagcatcac 7380
 cgtaaaaagg cccggttgc ggcgttttc aacccgacag gactataaag ataccaggcg 7440
 aaaaatcgac gtciaagtca gaggtggcga cctgttccga ccctgcccgt taccgatac 7500
 tttccccctg gaagctccct cgtgcgtct gcgcttctc aatgcctacg ctgttaggtat 7560
 ctgtccgcct ttctccctc gggaaagcgtg ctgggttgc tgacgacaacc ccccggttc 7620
 ctcagttcgg ttttaggtcgt tcgcgttcaag cgtcttgcgatc aaccccccgtt 7680
 cccgaccgct ggcgttccatc cggtaactat cgttttgcgatc aagacacgac 7740
 ttatcgccac tggcagcgc cactggtaac agatttagca gagcggaggt ttttaggtt 7800
 gctacagagt ttttgcgttcaac atctgcgtc tgctgaagcc agttacctc tacggctaca 7860
 aaacaaacca ccgtggtag cgggtttt gggatggat ttttttgcgatc agcggcggat 7920
 aaaaaaggat ctcaagaaga tcccttgcgatc ttttctacgg ggtctgacgc tcagtgaa 7980
 gaaaactcac gtttgcgttcaatc aatctaaatgaa ttttgcgttcaatc aaggatctt 8040
 cttttaaatt aaaaatgaa ttttgcgttcaatc gacagttacc aatgcgttcaatc atctaaatgaa 8100
 gacagttacc aatgcgttcaatc ctttttgcgatc ttttctacgg ggtctgacgc tcagtgaa 8160
 tccatagttt cctgactccc cgtcggttag ccacgttccatc aatctaaatgaa 8220
 gggcccgatg ctgcaatgtt accgcgagac aatctaaatgaa ttttgcgttcaatc aaggatctt 8280
 ataaaccagc cagccggaa ggcggagcgc aatctaaatgaa ttttgcgttcaatc aatctaaatgaa 8340
 atccagtcta ttaatttttcccgccggaaatc ttttctacgg ggtctgacgc tcagtgaa 8400
 cgcaacgtt ccgttccatc aatctaaatgaa ttttctacgg ggtctgacgc tcagtgaa 8460
 tcattcagct cccgttccca acgttccatc aatctaaatgaa ttttctacgg ggtctgacgc tcagtgaa 8520
 aaagcggtt gtccttcccg ccgttccatc aatctaaatgaa ttttctacgg ggtctgacgc tcagtgaa 8580
 tcactcatgg ttatggcaggc actgcataat ttttctacgg ggtctgacgc tcagtgaa 8640
 ttttctgtga ctgggtgagta ctcaaccaag ttttctacgg ggtctgacgc tcagtgaa 8700
 agttgtctt gcccggcgtt aacacgggat ttttctacgg ggtctgacgc tcagtgaa 8760
 gtgtcatca ttggaaaacg ttttctacgg ggtctgacgc tcagtgaa 8820
 agatccagtt cgatgttaacc cactcggtca cccaaactgtt ttttctacgg ggtctgacgc tcagtgaa 8880
 accagcggtt ctgggtgagc aaaaacagga aggcaaaatg cccggccggaaatg 8940
 ggcacacgga aatgttgaat actcataactc ttttctacgg ggtctgacgc tcagtgaa 9000
 cagggttatt gtctcatgag cggatacata ttttctacgg ggtctgacgc tcagtgaa 9060
 ggggttccgc gcacatttcc cccaaactgtt ttttctacgg ggtctgacgc tcagtgaa 9120
 atgacattaa cctataaaaa taggcgttac acgaggccct ttttctacgg ggtctgacgc tcagtgaa 9180
 ctcatgttttgcgttac acgttacatc ttttctacgg ggtctgacgc tcagtgaa 9240
 agatcgggaa cactgaaaaa taacagttt ttttctacgg ggtctgacgc tcagtgaa 9276

<210> 14

<211> 11

<212> PRT

<213> Pichia pastoris

<400> 14

Lys	Arg	Ile	Ala	Ser	Gly	Val	Ile	Ala	Pro	Lys
1										
						5				10

<210> 15

<211> 30

<212> DNA

```

<213> Artificial Sequence

<220>
<223> Primer

<400> 15
tgggtcttgg cacctgccag ctgctttctg 30

<210> 16
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 16
gaagggggaa gtggtgctgg gaccctag 28

<210> 17
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 17
ccctctgggt agccagcaca cagcatc 27

<210> 18
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 18
gccatcgtgg tgccggccaa ctacag 26

<210> 19
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 19
atcggtgggc cggccaaacta cagccaaatg 30

<210> 20
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

```

<400> 20
 accccatcacc tgctcccgta tccatgcctc 30

<210> 21
 <211> 103
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Polypeptide

<400> 21
 Val Ser Arg Leu Val Ile Ser Ile Arg Leu Pro Gln His Leu Gly Leu
 1 5 10 15
 Arg Pro Pro Leu Ala Leu Leu Glu Leu Ser Ser Arg Val Glu Pro Ser
 20 25 30
 Pro Ser Ala Leu Pro Ile Cys Leu His Pro Ala Gly Ile Pro Pro Gly
 35 40 45
 Ala Ser Cys Trp Val Leu Gly Trp Lys Glu Pro Gln Asp Arg Val Pro
 50 55 60
 Val Ala Ala Ala Val Ser Ile Leu Thr Gln Arg Ile Cys Asp Cys Leu
 65 70 75 80
 Tyr Gln Gly Ile Leu Pro Pro Gly Thr Leu Cys Val Leu Tyr Ala Glu
 85 90 95
 Gly Gln Glu Asn Arg Cys Glu
 100

<210> 22
 <211> 37
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Polypeptide

<400> 22
 Asn Asp Ser Arg Trp Ser Leu Leu Cys Gln Glu Glu Gly Thr Trp Phe
 1 5 10 15
 Leu Ala Gly Ile Arg Asp Phe Pro Ser Gly Cys Leu Arg Pro Arg Ala
 20 25 30
 Phe Phe Pro Leu Gln
 35